IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
HARLEY R. MYLER ET AL.

Serial No. **09/911,575**

Filing Date: **July 25, 2001**

For: **Method for Monitoring and Automatically Correcting Digital Video Quality By Reverse Frame Prediction**

Examiner: **Trang U. Tran**

Group Art Unit: **2614**

Attorney Docket No. **32707**

Asst. Commissioner for Patents
Alexandria, VA 22313-1450

Sir:

DECLARATION OF HARLEY R. MYLER UNDER 37 CFR § 1.132

I, Harley R. Myler, do hereby declare and say as follows:

1. I am a co-inventor in the above-referenced application along with Michele Van Dyke-Lewis for a method for monitoring and automatically correcting digital video quality by *Reverse Frame Prediction* (RFP). This application is currently under review by the USPTO. This Declaration in intended to present factual information and support for clarifying aspects of the invention as claimed. I am presently Chairman of the Department of Electrical Engineering at the Lamar University in Beaumont, Texas. I am the inaugural holder of the Mitchell Endowed Chair in Electrical Engineering at the rank of fully tenured full professor. I have both a PhD and MSEE from the New Mexico State University in Las Cruces, New Mexico and a BS in Chemistry with a double major in

electrical engineering from the Virginia Military Institute in Lexington, Virginia. I have published four books, three of which deal with image processing, along with a large number of refereed journal papers and conference proceedings.

2. The determination of digital video quality is a topic of intense debate and research by many top scientists and engineers around the globe. There is much degradation that can occur to digital video as it is passed through various filters and channels *en route* to a final destination. The Video Quality Experts Group (VQEG), an *ad hoc* organization with ties to the International Telecommunications Union (ITU), has defined three fundamental modes of digital video quality (VQ) analysis. These are the full-reference (FR), reduced-reference (RR) and the no-reference (NR). I coined the terminology "No Reference" at the March 2000 meeting of VQEG in Ottawa. At that time, many VQEG members were incredulous that a *no-reference* technique was possible. A FR technique simply compares an undegraded video stream to a potentially degraded stream and attempts to determine the disparity between the two against a human-derived subjective scale that has been coded into an algorithm. The nature of how the disparity is measured and how the human subjective measures are encoded is the subject of a volume of research and is ongoing at this time. A major disadvantage to the FR methods is the requirement that one have available the undegraded video at the time of analysis. Clearly, this is only useful in a laboratory setting and has no utility when video must be analyzed for quality at the terminus of a channel where the original stream is not available.

3. The RR methods attempt to make VQ analysis feasible at point-of-use, say, for example, in a home at the satellite-box receiving digital video. There are two fundamental approaches to RR, but both require an analysis of the undegraded video at the source. The analysis involves the extraction of specialized data features that attempt to describe the video in such a way that the quality aspects of the stream are extracted. A fundamental assumption is that the features describing the quality are of substantially less bandwidth, or data volume, than the source video. The features are then either *embedded* in the original video data or they are sent along with the video on

a reduced bandwidth channel. If the feature data is embedded, it suffers the possibility of being degraded by the channel that potentially degrades the video carrier. It can also, in itself, degrade the video. In the second approach, the cost of the second, reduced bandwidth channel is the major disadvantage to the technique.

4. Methods described in the present application fall into the NR classification in that the technique requires no reference data outside of the original digital video stream. No analysis is performed at the source and no secondary data channel is required. Other NR schemes do exist, but they include methods that look for specific degradations in the video that are expected to arise given the nature of the encoding and decoding algorithm used to process the video. These methods are not true VQ analyzers in that they are limited to specific *codecs* and to a pre-selected set of degradations. The present invention takes advantage of the large volume of data that is available in a digital video stream and the low probability of potential VQ altering degradations. Embodiments of the present invention do not process noise unless the noise occurs as a video degradation, as called for in Claim 1 by way of example, including producing at least one video quality metric, a difference from a noise metric. As will be addressed later in this declaration with reference to cited prior art, there is a distinction between channel induced noise, which is addressed by the codec employed, and video noise that is due to other causes, such as up or down sampling.

5. In September of this year my co-inventor and I, along with Dr. Jing Guo of McNeese State University in Lake Charles, Louisiana, published a paper in volume 50, issue 3 of the IEEE Transactions on Broadcasting entitled "Gabor Difference Analysis of Digital Video Quality." This paper discusses and presents an analysis of the Gabor difference technique for measuring video quality that is described in U.S. Patent No. 6,577,764 to Myler and Van Dyke-Lewis for a "Method for Measuring and Analyzing Digital Video Quality." This is a FR approach, but in the paper we also discuss RFP. The Transactions on Broadcasting is a refereed journal of the IEEE and publication of a paper in it requires the approval of three reviewers who are recognized experts in the area that the paper address and who are anonymous to the authors. During the review

of our paper, one reviewer commented:

> Quality: The paper addresses an issue of fundamental importance to the broadcasting industry. The overall quality of the paper is fair, its stronger points being relevance and originality and its weaker points being coherence, presentation and understandability.
>
> Originality: The paper is fairly original in the sense that there are no other methods of video quality assessment known to me that use the methodology advocated by the authors.
>
> Analysis: The paper doesn't appear to offer any concrete analysis in relation to the proposed methodology. Nevertheless this is not uncommon in papers dealing with this topic.

Such a review clearly supports the notion of originality and non-obviousness with regard to the RFP methods, as claimed, under any circumstance, even for those substantially skilled in the art.

6. To further support the argument that RFP is not obvious to one of ordinary skill in the art, the following is a transcript of an email that I received from Stephen Wolf on November 24, 1999 concerning the subject of "zero reference."

> *As to the checksum, I did say "checksum of sorts". I agree that the classic notion of checksum is such that it catches gross and common errors in transmitted data sets and could cause problems if not designed well. The idea of making it a perceptual checksum is good and that should be derivable from the subjective evaluation studies given a correct set of perceptually related features. But if you have a feature that measures a percept, could that feature then be used by itself? It should not be difficult to determine if the feature is proportional to scene degradation--of course, we are back to the universality of the feature. So what you are doing us saying "if this feature for this scene is greater than X, then the scene is degraded," where X is supplied with the frame. How far have you gotten with this idea?*
> ***I wish I could tell you that we have solved this problem, but we have not even tried to solve the zero reference problem.*** *The closest thing we have done to the content analyzer you mentioned is to classify scenes according to their spatial and temporal information content on a two dimensional grid. If you*

download itu9435e.pdf at http://www.its.bldrdoc.gov/n3/video/documents.htm, you will see this done for the ANSI test scenes.

S.W.

Stephen Wolf is the Project Leader of the Video Quality Research Program at the Institute for Telecommunication Sciences (ITS), an office of the National Telecommunications and Information Administration (NTIA) in Boulder, Colorado. He has published numerous papers and holds three U.S. patents for the development RR video quality measurement systems. Zero reference is an alternate terminology for NR.

7. I have read the Office Action dated May 18, 2004, including the Examiner's opinions regarding patentability of claims in the case. Further, I have read and am familiar with the references cited by the Examiner including U.S. Patent No. 5,767,922 to Zabih et al., U.S. Patent No. 5,745,169 to Murphy et al., and U.S. Patent No. 5, 261,030 to Tanaka used to reject claims in view of what the Examiner believes to be taught by applicants admitted prior art (AAPA) presented in the specification. I believe that the Examiner has incorrectly used these cited references in rejecting claims to a method and system for correcting errors in digital video as presented in the above referenced application

8. While these cited references appear to offer various elements of RFP, there is no motivation to combine these elements given the evidence described above. By way of supporting example, consider the teachings of the references hereafter referred to as Murphy '169, Zabih '922, and Tanaka '030. By way of explanation, Murphy '169 is directed to an approach to error detection that is fundamentally different from quality analysis. This patent teaches the use of "blocks of frequency related coefficients representing respective blocks of the image" and the analysis of "the statistical distribution of said coefficients within a block to detect the presence of corrupted data." Features in Murphy '169 lie with the intent to detect data corruption as a result of errors in coded coefficients. These coefficients represent a video frame reduced to coded blocks, not the entire video frame being processed for VQ degradation as presented by

the teachings of the present invention. Although Murphy '169 analyzes block difference statistics, one of skill in the art would not look to Murphy '169, nor would it be obvious to one of skill in the art, to extend this to frame-to-frame comparisons.

9. By way of further example, consider the language of claim 1 in the present application. Such language does not induce the granularity of a blockwise analysis of the suspect frame. In particular, the use of the *at least one intercut sequence* insures that the collective statistic is based on a frame-wise analysis and not blocks. While calling for an intercut sequence may appear to evoke the teachings of Zabih '922, which specifies "A process for detecting scene breaks in a sequence of images" by "detecting the location of intensity edges in every image and storing the information about the location of the intensity edges in a plurality of successive binary images," there is a clear distinction. In this case, the teachings of Zabih '922 include the detection of scene breaks and the methodology includes the use of intensity images. As illustrated by way of example with reference to Claim 8, the present invention calls for scene breaks, determined from *interframe correlation* and not edge detection, to establish a set of frames for analysis. This again is not an obvious construction and there is no motivation for the combination of the Murphy '169 approach to error detection with the Zabih '922 scene break method to produce RFP.

10. What is significant in Tanaka '030 is the motivation for *motion compensation* (MC) analysis that is a distinctive motivation opposed to the frame-to-frame collective analysis methods of VQ. Although Tanaka '030 calls for a "predicting apparatus comprising means for outputting an input block by forming a block from a digital image signal stream within a frame," there is again no motivation to combine this with Murphy '169 and Zabih '922 in order to produce the VQ analysis taught by our *Reverse Frame Prediction* technique. Tanaka '030 is looking to account for motion and reduce the errors generated by motion in the video sequence, not analyze a set of frames for VQ.

11. While a collection of known prior art including Murphy '169, Zabih '922, and

Tanaka '030 may teach, in various ways, elements of RFP, there is no clear evidence of a motivation to combine given the evidence that we submitted above showing that our approach is entirely novel to experts in video quality analysis.

12. I hereby declare that all statements made herein of my own accord are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that any such willful false statements may jeopardize the validity of the application or any patent issued thereon.

11.16.4
Date

Harley R. Myler